服务器部署

服务器系统：ubuntu 16.04

Elasticsearch版本：5.1

开发语言：python

web框架：Django

web服务器：apache

爬虫框架：Scrapy

**ES部署：**

apt-get install python-software-properties

apt-get install software-properties-common

添加jdk 1.8 环境

添加oracle的源：

sudo add-apt-repository ppa:webupd8team/java

sudo apt-get update

安装

sudo apt-get install oracle-java8-installer

安装python-pip和你所需要的包：

apt-get install python-pip

检查你所安装Pip的版本：

pip-V

安装git：

sudo apt-get install git

下载es

git clone git://github.com/medcl/elasticsearch-rtf.git -b master --depth 1

**apache部署：**

安装各种软件**：**

apache sudo apt-get install apache2 安装后请使用apachectl -v来检查版本号 2.4.x与2.2.x后续有一点区别

django sudo pip install Django==1.8

* 建立Django与Apache的连接

sudo apt-get install libapache2-mod-wsgi

**然后新建一个网站的配置文件**

* sudo vi /etc/apache2/sites-available/yoursite.conf
* 配置文件的具体内容如下 `

ServerName www.yourdomain.com # 改为你自己的域名

# ServerAlias otherdomain.com

# ServerAdmin youremail@gmail.com

# 存放用户上传图片等文件的位置，注意去掉#号

#Alias /media/ /var/www/ProjectName/media/

# 静态文件(js/css/images)

Alias /static/ /var/www/ProjectName/static/

# 允许通过网络获取static的内容

<Directory /var/www/ProjectName/static/>

Require all granted

</Directory>

# 最重要的！通过wsgi.py让Apache识别这是一个Django工程，别漏掉前边的 /

WSGIScriptAlias / /var/www/ProjectName/ProjectName/wsgi.py

# wsgi.py文件的父级目录，第一个ProjectName为Django工程目录

# 第二个ProjectName为Django自建的与工程同名的目录

<Directory /var/www/ProjectName/ProjectName/>

<Files wsgi.py>

Require all granted

</Files>

</Directory>

</VirtualHost>`

#### **需要注意的是 假如你的Apache版本为2.2.x 则将Require all granted改为Order deny,allow Allow from all**

* 然后 执行sudo a2ensite yoursite.conf来使网站生效
* 也可以执行sudo a2dissite yoursite.conf来使网站失效
* 最后重启Apache即可 sudo service apache2 restart
* e.修改Django的wsgi.py文件
* 路径/var/www/ProjectName/ProjectName/wsgi.py
* 修改为以下内容：
* ```import os from os.path import join,dirname,abspath PROJECT\_DIR = dirname(dirname(abspath(file)))

import sys sys.path.insert(0,PROJECT\_DIR) os.environ.setdefault("DJANGO\_SETTINGS\_MODULE", "project.settings")

from django.core.wsgi import get\_wsgi\_application application = get\_wsgi\_application()

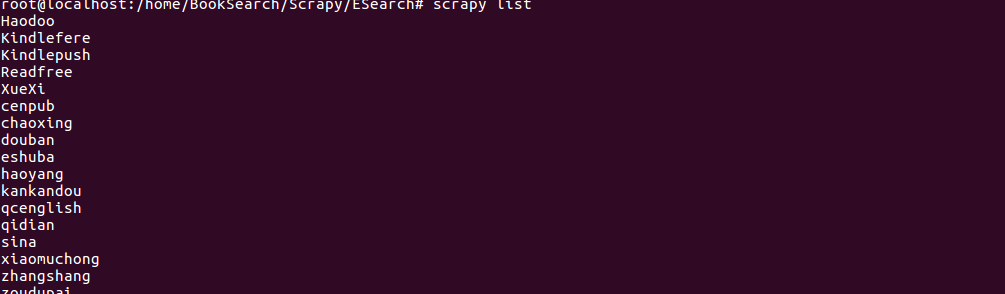
* 注意将"project.settings" 改为正确的名称
* 再次重启Apache sudo service apache2 restart

scrapy部署：

scrapyd安装： pip install scrapyd

pip install scrapyd-client

scrapy list 查看爬虫个数：

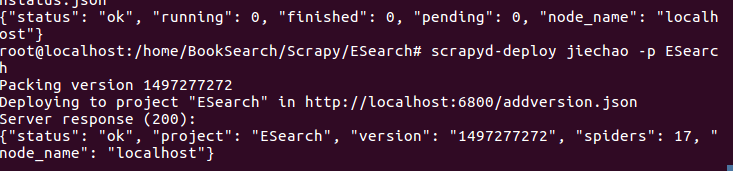


scrapyd使用 ：在scrapy文件目录下运行nohup scrapyd & 后台运行

scrapyd-client使用 :

项目上传：在scrapy文件下运行 scrapyd-deploy jiechao -p ESearch

打印信息如下：



项目上传完成！

运行爬虫：

curl http://localhost:6800/schedule.json -d project=myproject -d spider=somespider

